GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 12:37:33 ; Search time 189 Seconds
                                         (without alignments)
                                         769.494 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCCIASPSYPEI 331
Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5
Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_21:*
             1: geneseqp1908:*
             2: geneseqp1908:*
             3: geneseqp2008:*
             4: geneseqp2015:*
             5: geneseqp2025:*
             6: geneseqp2003ss:*
             7: geneseqp2003bs:*
             8: geneseqp2004s:*
             9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   %
No.   Score   Query   Match   Length   DB   ID   Description
               ------   -----   ------   --   --   -----------
 1   1747   97.1   331   7   ADP65168   Adp65168 Human cat
 2   1747   97.1   331   9   ADY14944   Ady14944 PRO polyp
 3   1747   97.1   331   9   ADY14946   Ady14946 PRO polyp
 4   1747   97.1   331   9   AEC01587   Aec01587 Human cat
 5   1744   96.9   331   7   ADD45416   Add45416 Human Pro
 6   1744   96.9   331   7   ADE56902   Ade56902 Human Pro
 7   1744   96.9   331   8   ADJ71695   Adj71695 Human NOV
 8   1741   96.8   331   7   ADD35927   Add35927 Human cat
 9   1740   96.7   331   8   AAY59634   Aay59634 Human Cat
10   1738   96.6   331   9   ADY19784   Ady19784 PRO polyp
11   1734   96.4   331   7   ADD35932   Add35932 Human cat
12   1731   96.2   331   7   ADD35934   Add35934 Human cat
13   1724   95.6   331   7   ADD35933   Add35933 Human cat
14   1701.5 94.6   330   6   AAO30451   Aao30451 Monkey ca
15   1631   90.7   330   6   AAO30451   Aao30451 Monkey ca
16   1631   90.7   313   8   ABM83324   Abm83324 Human dia
17   1631   90.7   313   8   ABM83323   Abm83323 Human dia
18   1600   88.9   331   7   ABR56160   Abr56160 Canine ca
19   1580   87.8   331   6   ABG73437   Abg73437 Human Cat
20   1388   77.2   267   6   ADG69537   Adg69537 Human hea
21   1309.5 72.8   330   7   ADD45414   Add45414 Rat Prote
22   1309.5 72.8   330   7   ADE56900   Ade56900 Rat Prote
23   1145   63.6   217   7   ADD35931   Add35931 Human mat
24   1144   63.6   222   8   ADJ71697   Adj71697 Human NOV
25   1142   63.5   217   9   ADY58885   Ady58885 Human cat

RESULT 1
ID   ADP65168
AC   ADP65168;
XX
AC   ADP65168 standard; protein; 331 AA.
XX
DT   12-AUG-2004 (first entry)
XX
DE   Human cathepsin S preproprotein.
XX
PN   WO200307282 7-A1.
XX
PD   04-SEP-2003.
XX
PF   31-OCT-2002; 2002WO-US035433.
PR   31-OCT-2001; 2001US-0336220P.
XX
PA   (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PI   Hirsch R, Thorton SL;
XX
DR   WPI; 2003-712740/67.
DR   GENBANK; NP_004070.
XX
PT   Diagnosing and analyzing autoimmune disease using gene expression
PT   profiles and microarray technology, useful for diagnosing and treating
PT   rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT   gout.
XX
PS   Disclosure; Page; 56pp; English.
XX
CC   The invention relates to a novel method for diagnosing and analysing
CC   autoimmune disease or arthritides. The method comprises obtaining a
CC   patient sample containing mRNA, analysing gene expression using the mRNA
CC   that results in a gene expression signature of the mRNA, and using that
CC   gene expression signature to diagnose or analyse the autoimmune disease
CC   or arthritides in the patient, where gene expression of at least 60% of

                              ALIGNMENTS

25   1138   63.3   217   7   ADD35928   Add35928 Human mat
26   1135   63.1   217   7   ADD35930   Add35930 Human mat
27   1128   62.7   217   7   ADD35929   Add35929 Human mat
28   963   53.5   330   9   ADZ11834   Adz11834 Canine ca
29   960.5 53.4   329   2   AAR82720   Aar82720 Human ost
30   960.5 53.4   329   2   AAR95599   Aar95599 Prepro-ca
31   960.5 53.4   329   2   AAW41645   Aaw41645 Human cat
32   960.5 53.4   329   3   AAW39216   Aaw39216 Human cat
33   960.5 53.4   329   3   AAY81000   Aay81000 Human pro
34   960.5 53.4   329   7   ADE60397   Ade60397 Human Pro
35   960.5 53.4   329   8   ADD45861   Add45861 Protein e
36   960.5 53.4   329   8   ADQ18392   Adq18392 Human sof
37   960.5 53.4   329   8   ABM80141   Abm80141 Tumour-as
38   960.5 53.4   329   8   ADP12997   Adp12997 Human Pro
39   960.5 53.4   329   9   ADP25400   Adp25400 PRO polyp
40   960.5 53.4   329   9   AEC01585   Aec01585 Human cat
41   960.5 53.4   361   3   AAB57115   Aab57115 Human pro
42   958.5 53.3   329   3   AAB30506   Aab30506 A monkey
43   956.5 53.2   329   2   AAW77071   Aaw77071 Rat Cathe
44   954   53.0   329   2   ADE60395   Ade60395 Rat Prote
45   954   53.0   329   7   ADD45859   Add45859 Rat Prote

ALIGNMENTS

ADP65168

Homo sapiens.

OS
XX
KW   immune; human.
KW   fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW   immunomodulatory; lupus; ankylosing spondylitis; fibrositis;
KW   antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW   rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW   autoimmune disease; arthritide; gene expression analysis;
DE   Human cathepsin S preproprotein.

```
OM protein - protein search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             November 21, 2005, 12:42:19 ; Search time 232 Seconds
                                               (without alignments)
                                               1006.594 Million cell updates/sec

Title:              US-10-809-816A-2
Perfect Score:      1799
Sequence:           1 MKQLVCVLLVCSSAVAQLHK.........MARNKGNHCGIASPSYPEI 331

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:           2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters:      2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          UniProt_05.80:*
                    1: uniprot_sprot:*
                    2: uniprot_trembl:*

                    SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result          Query
No.   Score   Match  Length  DB  ID            Description
                *
1     1747    97.1           2   Q5T5I0_HUMAN  homo sapien
2     1744    96.9           2   CATS_HUMAN    homo sapien
3     1744    96.9           2   Q6FH55_HUMAN  homo sapien
4     1701.5  94.6           1   CATS_SAIBB    saimiri bol
5     1600    88.9           1   CATS_CANFA    canis famil
6     1406.5  78.2           2   Q6PCU5_RAT    rattus norv
7     1378.5  76.6           2   Q8BSZ5_MOUSE  mus musculu
8     1377.5  76.0           2   CATS_MOUSE    mus musculu
9     1366.5  76.0           2   Q99M14_MOUSE  mus musculu
10    1309.5  72.8           1   CATS_RAT      rattus norv
11    1230.5  68.4           2   Q5ZMR6_CHICK  gallus gall
12    1140    63.4           2   Q6DJC1_XENTR  xenopus tro
13    1136.5  63.2           2   Q6DE57_XENLA  xenopus lae
14    1112.5  61.8           3   Q7T0S4_XENLA  xenopus lae
15    1094    60.8           2   Q7T183_9ICH   paralabidoc
16    1070    59.5           2   Q502A6_BRARE  brachydanio
17    1065    59.2           2   Q502H6_BRARE  brachydanio
18    1050    58.4           2   Q566T8_BRARE  brachydanio
19    1047    58.2           2   Q4QRH8_BRARE  brachydanio
20    1038.5  57.7           2   Q58HF5_PAROL  paralichthy
21    1036    57.6           1   CATS_BOVIN    bos taurus
22    1026    57.0           2   Q6J2Z4_FUNHE  fundulus he
23    1022    56.8           1   Q90324_CYPCA  cyprinus ca
24    1004    55.8           2   Q4SW28_TETNG  tetraodon n
25    983.5   54.7           2   Q6F6A0_ORYLA  oryzias lat
26    968.5   53.8           1   CATK_PIG      sus scrofa
27    968.5   53.8           1   CATK_BOVIN    bos taurus
28    965.5   53.7           1   Q5E968_BOVIN  bos taurus
29    961.5   53.4           1   CATK_RABIT    oryctolagus
30    961.5   53.4           1   CATK_MACFA    macaca fasc
31    960.5   53.4           1   CATK_HUMAN    homo sapien
32    960.5   53.4           2   Q6FHN2_HUMAN  homo sapien
33    960.5   53.4           2   Q6FH56_HUMAN  homo sapien
34    954     53.0           1   CATK_RAT      rattus norv
35    939     52.2           2   O35186        rattus norv
36    937     52.1           2   Q6PA77_XENLA  xenopus lae
37    935     52.0           2   Q6DJA7_XENTR  xenopus tro
38    934     51.9           1   Q5BL90_XENTR  xenopus tro
39    934     51.9           2   CATK_MOUSE    mus musculu
40    927.5   51.6           2   Q54570_MOUSE  mus musculu
41    926     51.5           2   Q6XRF7_FUNHE  fundulus he
42    919     51.1           2   Q75SZ8_CYPCA  cyprinus ca
43    914     50.8           2   Q6F6A1_ORYLA  oryzias lat
44    912.5   50.7           2   Q4SW27_TETNG  tetraodon n
45    912     50.7           2   Q68F41_XENLA  xenopus lae
                               2   Q90WC2_ONCMY oncorhynchu
```

ALIGNMENTS

```
RESULT 1
Q5T5I0 HUMAN
ID   Q5T5I0_HUMAN PRELIMINARY;     PRT;   331 AA.
AC   Q5T5I0;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Cathepsin S.
GN   Name=CTSS; ORFNames=RP11-363I22.2-001;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Howden P.;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL356292; CAI13657.1; -; Genomic_DNA.
DR   GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
SQ   SEQUENCE   331 AA;  37495 MW;  86093619DB6F0269 CRC64;

Query Match          97.1%;    Score 1747;  DB 2;  Length 331;
Best Local Similarity   97.0%;  Pred. No. 2.5e-135;
Matches  321;  Conservative  2;  Mismatches  8;  Indels  0;  Gaps  0;

Qy        1  MKQLVCVLLVCSSAVAQLHKDPTLDHHWHLWKKTYGKQYKEKNEEAVRRLIWEKNLKFVM  60
             ||:|||||||||||||||||||||| | |||||||||||||||||||||||||||||||
Db        1  MKRLVCVLLVCSSAVAQLHKDPTLDHHWHLWKKTYGKQYKEKNEEAVRRLIWEKNLKFVM  60

Qy       61  LHNLEHSMGMHSYDLGMNHLGDMTSEEVMSLMSSLRVPSQWQRNITYKSNANQILPDSVD  120
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       61  LHNLEHSMGMHSYDLGMNHLGDMTSEEVMSLMSSLRVPSQWQRNITYKSNPNRILPDSVD  120

Qy      121  WREKGCVTEVKYQGSCGACWAFSAVGALEAQLKLKTGKLVSLSAQNLVDCSTEKYGNKGC  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121  WREKGCVTEVKYQGSCGACWAFSAVGALEAQLKLKTGKLVSLSAQNLVDCSTEKYGNKGC  180

Qy      181  NGGFMTRAFQYIIDNNGIDSDASYPYKATDQKCQYDSKYRAATCSKYTELPYGREDVLKE  240
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      181  NGGFMTAFQYIIDNKGIDSDASYPYKAMDQKCQYDSKYRAATCSKYTELPYGREDVLKE  240

Qy      241  AVANKGPVSVGVDARHPSFFLYRSGVYYEPSCTQNVNHGVLVVGYGVLNGKEYWLVKNSW  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241  AVANKGPVSVGVDARHPSFFLYRSGVYYEPSCTQNVNHGVLVVGYGVLNGKEYWLVKNSW  300

Qy      301  GRNFGEEGYIRMARNKGNHCGIASPPSYPEI 331
             ||||||||||| | ||||||||||||||||
Db      301  GHNFGEEGYIRMARNKGNHCGIASPSYPEI 331

RESULT 2
CATS_HUMAN
```

OM protein - protein search, using sw model

GenCore version 5.1.6

Run on: November 21, 2005, 12:45:19 ; Search time 41 Seconds
(without alignments)
776.774 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCGIASPFSYPEI 331
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1738	96.6	331	2	A42482	cathepsin S (EC 3.
2	1309.5	72.8	330	2	A44087	cathepsin S (EC 3.
3	1036	57.6	217	2	S15844	cathepsin S (EC 3.
4	965.5	53.7	329	2	A49868	cathepsin K (EC 3.
5	960.5	53.4	329	2	JC2476	cathepsin K (EC 3.
6	930.5	51.7	329	2	S74427	cathepsin K (EC 3.
7	862	47.9	334	1	KHMSL	cathepsin L
8	862	47.9	334	1	KHRTL	cathepsin L (EC 3.
9	848.5	47.2	333	1	KHHUL	cathepsin L (EC 3.
10	837	46.5	334	1	A58195	cathepsin L (EC 3.
11	792	44.0	313	2	A58195	cathepsin L (EC 3.
12	786	43.7	323	2	S19650	cysteine proteinas
13	786	43.7	339	2	A53810	cathepsin L (EC 3.
14	779.5	43.3	344	2	JX0366	cysteine endopepti
15	778.5	43.3	326	2	S53027	cathepsin L (EC 3.
16	777.5	43.2	338	2	JC5443	cathepsin L-like c
17	774.5	43.1	337	2	T24387	probable cysteine
18	770	42.8	320	2	S19651	cysteine proteinas
19	751.5	41.8	333	1	I52525	testin precursor -
20	743	41.3	343	2	JC7183	cathepsin Q (EC 3.
21	740.5	41.2	331	2	JC5441	cathepsin L-like c
22	737.5	41.0	324	2	S47432	cathepsin L (EC 3.
23	737.5	41.0	331	2	JC5542	cysteine proteinas
24	735.5	40.9	322	2	S19649	cysteine proteinas
25	724	40.2	326	2	S43991	cathepsin L-like p
26	704.5	39.2	218	2	KHCHL	cathepsin L (EC 3.
27	699	38.9	480	2	T01207	cysteine proteinas
28	690	38.4	326	2	T09929	cathepsin L-like p
29	689	38.3	458	1	KHRZOA	oryzain (EC 3.4.22
30	672.5	37.4	218	2	S67481	cathepsin L-like c
31	661	36.7	462	2	JN0719	drought-inducible
32	649	36.1	427	2	S57776	cysteine proteinas
33	647.5	36.0	455	2	T12041	cysteine proteinas
34	647	36.0	317	2	S44151	cathepsin L (EC 3.
35	635	35.3	368	2	S47312	cysteine proteinas
36	630.5	35.0	331	2	D86413	cysteine proteinas
37	629.5	35.0	454	2	JC4848	cysteine proteinas
38	624	34.7	367	2	T06529	cysteine proteinas
39	622.5	34.6	374	2	T03941	cysteine proteinas
40	618.5	34.4	464	2	S24602	oryzain (EC 3.4.22
41	618.5	34.4	471	1	KHRZOB	oryzain (EC 3.4.22
42	618	34.4	355	2	T05390	probable cysteine
43	617.5	34.3	376	1	KHDOP	prestalk cathepsin
44	615	34.2	380	1	TAGB	actinidain (EC 3.4
45	614.5	34.2	416	2	G86232	cysteine proteinas

ALIGNMENTS

RESULT 1
A42482
cathepsin S (EC 3.4.22.27) precursor - human
C;Species: Homo sapiens (man)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42482; A53625; A42896
R;Shi, G.P.; Munger, J.S.; Meara, J.P.; Rich, D.H.; Chapman, H.A.
J. Biol. Chem. 267, 7258-7262, 1992
A;Title: Molecular cloning and expression of human alveolar macrophage cathepsin S, an e
A;Reference number: A42482; MUID:92218373; PMID:1373132
A;Accession: A42482
A;Molecule type: mRNA
A;Residues: 1-331 <SHI>
A;Cross-references: UNIPROT:P25774; UNIPARC:UPI00001517CF0; GB:U07374
A;Experimental source: alveolar macrophage
R;Wiederanders, B.; Bromme, D.; Kirschke, H.; von Figura, K.; Schmidt, B.; Peters, C.
J. Biol. Chem. 267, 13708-13713, 1992
A;Title: Phylogenetic conservation of cysteine proteinases. Cloning and expression of a
A;Reference number: A42896; MUID:92317106; PMID:1377692
A;Accession: A42896
A;Status: not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-91,'M',93-160,'S',162-331 <WIE>
A;Cross-references: UNIPARC:UPI000013DFE1; GB:S39127; GB:M90696; NID:g250802; PIDN:AAA09
A;Note: sequence extracted from NCBI backbone (NCBIP:107806)
R;Shi, G.P.; Webb, A.C.; Foster, K.E.; Knoll, J.H.M.; Lemere, C.A.; Munger, J.S.; Chapma
J. Biol. Chem. 269, 11530-11536, 1994
A;Title: Human cathepsin S: chromosomal localization, gene structure, and tissue distrib
A;Reference number: A53625; MUID:94209337; PMID:8157683
A;Accession: A53625
A;Molecule type: DNA
A;Residues: 1-210,'H',212-331 <SH2>
A;Cross-references: UNIPARC:UPI000013DEF1; GB:M90696; NID:g250802; PIDN:AAA09
A;Experimental source: testis
A;Map position: 1q21-1q21
A;Cross-references: GDB:132414; OMIM:116845
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase; lysosome
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-114/Domain: propeptide #status predicted <PRO>
F;115-331/Product: cathepsin S #status predicted <MAT>
F;139,278,298/Active site: Cys, His, Asn #status predicted

Qy

Query Match   96.6%; Score 1738; DB 2; Length 331;
Best Local Similarity   96.4%; Pred. No. 9.7e-136;
Matches 319; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy   1 MKQLVCVLLVCSSAVAQLHKDPTLDHHWHLWKKTYGKQYKENEAVRRLLWEKNLKFVM 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKQLVCVLLVCSSAVAQLHK

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 22, 2005, 11:13:59 ; Search time 806 seconds
(without alignments)
3395.985 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK.........MARNKGNHCGIASPSYPEI 331

Searched: 9793542 seqs, 413469005 residues

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Total number of hits satisfying chosen parameters: 19587084

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10809816/runat 21112005_094203_8378/app_query.fasta_1.519
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10809816 @CGN_1_1_1549 @runat 21112005_094203_8378 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:*
1: /cgn2_6/ptcdata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptcdata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptcdata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptcdata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptcdata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptcdata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptcdata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptcdata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptcdata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptcdata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1799	100.0	1798	9	US-10-809-816A-1	Sequence 1, Appli
2	1747	97.1	1784	7	US-10-641-643-1227	Sequence 1227, Ap
3	1747	97.1	4100	9	US-10-973-858-19	Sequence 19, Appl
4	1747	97.1	4100	9	US-10-287-436A-51	Sequence 51, Appl
5	1744	96.9	1254	8	US-10-635-398-95	Sequence 95, Appl
6	1744	96.9	1752	9	US-10-956-157-1177	Sequence 1177, Ap
7	1744	96.9	1752	9	US-10-287-436A-953	Sequence 953, App
8	1740	96.7	1643	5	US-10-099-275-1	Sequence 1, Appli
9	1740	96.7	1763	9	US-09-873-367C-173	Sequence 173, App
10	1740	96.7	1763	9	US-10-843-641A-173	Sequence 173, App
11	1740	96.7	1763	9	US-10-287-436A-954	Sequence 954, App
12	1738	96.6	1255	9	US-10-172-118-987	Sequence 987, App
13	1738	96.6	1255	7	US-10-342-987-987	Sequence 987, Appl
14	1738	96.6	1255	7	US-10-283-975A-78	Sequence 78, Appl
15	1738	96.6	1255	8	US-10-278-698-206	Sequence 206, Appl
16	1738	96.6	1255	5	US-10-278-698-722	Sequence 722, App
17	1701.5	94.6	993	5	US-10-010-577-1	Sequence 1, Appl
18	1701.5	94.6	993	7	US-10-894-104-1	Sequence 1, Appl
19	1680	93.4	3164	7	US-10-641-643-816	Sequence 816, App
20	1600	88.9	996	8	US-10-894-106-1	Sequence 1, Appl
21	1580	87.8	996	8	US-09-990-064-1	Sequence 1, Appl
22	1570	87.3	1126	5	US-10-010-580-1	Sequence 1, Appli
23	1318.5	73.3	1330	7	US-10-191-803-109	Sequence 109, App
24	1318.5	73.3	1330	9	US-10-152-319A-1700	Sequence 1700, Ap
25	1318.5	73.3	1330	9	US-10-486-706-131	Sequence 131, App
26	1239	68.9	1400	9	US-10-956-157-6412	Sequence 6412, Ap
27	1144	63.6	667	9	US-10-635-398-97	Sequence 97, Appl
28	963	53.5	1001	9	US-10-929-919A-31	Sequence 31, Appl
29	963	53.5	1001	9	US-10-318-584-1	Sequence 1, Appl
30	960.5	53.4	1482	6	US-10-723-860-1210	Sequence 1210, Ap
31	960.5	53.4	1482	9	US-10-756-149-1175	Sequence 1175, Ap
32	960.5	53.4	1482	8	US-09-953-956-1	Sequence 1, Appl
33	960.5	53.4	1619	9	US-10-114-464-1	Sequence 1, Appli
34	960.5	53.4	1619	8	US-10-726-645-1	Sequence 1, Appli
35	960.5	53.4	1619	6	US-10-172-118-479	Sequence 479, App
36	960.5	53.4	1669	6	US-10-240-425-1556	Sequence 1556, Ap
37	960.5	53.4	1669	7	US-10-342-987-479	Sequence 479, App
38	960.5	53.4	1669	9	US-10-956-157-408	Sequence 408, Appl
39	960.5	53.4	1671	9	US-10-044-090-462	Sequence 462, Appl
40	960.5	53.4	1702	8	US-10-278-698-204	Sequence 204, Appl
41	960.5	53.4	1702	8	US-10-278-698-720	Sequence 720, App
42	960.5	53.4	1702	8	US-09-925-300-17	Sequence 17, Appl
43	960.5	53.4	1756	3	US-10-973-858-753	Sequence 753, App
44	960.5	53.4	2017	5	US-10-198-846-10186	Sequence 10186, A
45	960.5	53.4	2042	8	US-10-357-930-23382	Sequence 23382, A

ALIGNMENTS

RESULT 1
US-10-809-816A-1

Sequence 1, Application US/10809816A
Publication No. US20050214774A1
GENERAL INFORMATION:
APPLICANT: Li, Shyr-Jiann et al.
TITLE OF INVENTION: ISOLATED MONKEY CATHEPSIN S PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING MONKEY CATHEPSIN S PROTEINS,
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CL001507
CURRENT APPLICATION NUMBER: US/10/809,816A
CURRENT FILING DATE: 2004-03-26
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 1
LENGTH: 1798
TYPE: DNA
ORGANISM: Cynomologous monkey
US-10-809-816A-1

Alignment Scores: Length: 1798
Pred. No.: 2.17e-216   Matches: 331
Score: 1799.00   Conservative: 0
Percent Similarity: 100.00%   Mismatches: 0
Best Local Similarity: 100.00%   Indels: 0
Query Match: 9   Gaps: 0
DB:

US-10-809-816A-2 (1-331) x US-10-809-816A-1 (1-1798)

Qy 1 MetLysGlnLeuValCysValLeuLeuValCysSerSerAlaValAlaGlnLeuHisLys 20

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 22, 2005, 12:16:10 ; Search time 224 Seconds
(without alignments)
194.229 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCGIASPSYPEI 331

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3196817 seqs, 65720914 residues

Total number of hits satisfying chosen parameters: 6393634

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US10809816/runat_21112005_094204_8411/app_query.fasta_1.519
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN10 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10809816 @CGN_1_1.183 @runat_21112005_094204_8411
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	859	47.7	1547	1	US-10-131-826A-11	Sequence 11, Appl
	2	254.5	14.1	2197	1	US-10-131-826A-313	Sequence 313, App
c	3	81	4.5	3721	1	US-10-131-826A-543	Sequence 543, App
c	4	79.5	4.4	2760	9	US-11-008-331-1	Sequence 1, Appli
c	5	78.5	4.4	4040	1	US-10-131-826A-425	Sequence 425, Appl
c	6	75.5	4.2	4338	1	US-10-526-731-2	Sequence 2, Appl
c	7	75.5	4.2	9719	9	US-11-042-988-10	Sequence 10, Appl
	8	74.5	4.1	2745	1	US-10-647-956A-5	Sequence 5, Appli
	9	70.5	3.9	1777	1	US-10-632-150-51	Sequence 51, Appl
	10	70.5	3.9	2602	1	US-10-510-386-229	Sequence 229, Appl
	11	69.5	3.9	1089	1	US-10-660-499A-1	Sequence 1, Appl
	12	68.5	3.8	1299	1	US-10-689-742-143	Sequence 143, Appli
	13	68.5	3.8	1799	7	US-11-071-651-13	Sequence 13, Appli
	14	68.5	3.8	2584	1	US-10-510-386-101	Sequence 101, Appl
	15	68.5	3.8	8931	7	US-11-071-651-16	Sequence 16, Appl
c	16	67.5	3.8	2024	1	US-10-131-826A-259	Sequence 259, App
	17	67.5	3.8	3479	1	US-10-510-386-197	Sequence 197, Appl
	18	67	3.7	1677	1	US-10-131-826A-45	Sequence 45, Appl
	19	67	3.7	2237	7	US-11-102-978-3	Sequence 3, Appl
	20	67	3.7	340000	1	US-10-957-569-50	Sequence 50, Appli
	21	66.5	3.7	3048	1	US-10-510-386-27	Sequence 27, Appli
	22	66.5	3.7	3625	1	US-10-131-826A-123	Sequence 123, App
	23	66.5	3.7	4041	1	US-10-510-386-199	Sequence 199, Appl
	24	66	3.7	1368	7	US-11-074-176-49	Sequence 49, Appli
	25	66	3.7	1670	1	US-10-614-599-2	Sequence 2, Appli
c	26	65.5	3.6	993	9	US-11-082-389-415	Sequence 415, Appl
c	27	65.5	3.6	1209	9	US-11-082-389-417	Sequence 417, Appl
	28	65	3.6	645	1	US-11-074-176-359	Sequence 359, Appl
	29	65	3.6	675	7	US-11-074-176-233	Sequence 233, Appl
	30	65	3.6	2347	1	US-10-510-386-23	Sequence 23, Appl
	31	65	3.6	2859	1	US-10-689-742-99	Sequence 99, Appl
	32	64.5	3.6	693	1	US-10-131-826A-499	Sequence 499, Appl
	33	64	3.6	4158	1	US-10-971-560-3	Sequence 3, Appli
	34	64	3.6	6458	1	US-10-971-560-1	Sequence 1, Appli
	35	63.5	3.5	1675	1	US-10-131-826A-135	Sequence 135, Appli
	36	63.5	3.5	3283	1	US-10-131-826A-357	Sequence 357, App
	37	63	3.5	2388	9	US-11-082-389-99	Sequence 99, Appl
	38	63	3.5	2388	9	US-11-082-389-101	Sequence 101, Appl
	39	63	3.5	3639	7	US-11-074-176-255	Sequence 255, App
c	40	63	3.5	1095	7	US-11-082-389-439	Sequence 439, App
c	41	62.5	3.5	4374	1	US-10-131-826A-125	Sequence 125, Appl
c	42	62.5	3.5	6159	7	US-11-013-759-8	Sequence 8, Appli
c	43	62.5	3.5	13943	9	US-11-012-762-39	Sequence 39, Appli
c	44	62	3.4	1510	1	US-10-131-826A-399	Sequence 399, Appl
	45	62	3.4	2030	1	US-10-652-893-3	Sequence 3, Appli

RESULT 1
US-10-131-826A-11
; Sequence 11, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113

ALIGNMENTS

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 22, 2005, 09:48:49 ; Search time 205 Seconds
           (without alignments)
           2870.112 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCGIASFPSYPEI 331

Searched: 1303057 seqs, 888780828 residues

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10809816/runat_21112005_094202_8310/app_query.fasta_1.519
-DB=issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=10
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10809816 @CGN_1_1_290 @runat_21112005_094202_8310 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
         1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
         2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
         3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
         6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
         7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
         8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
         9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1747	97.1	1784	3	US-09-023-655-1227	Sequence 1227, Ap
2	1740	96.7	1643	3	US-09-701-685-1	Sequence 1, Appli
3	1701.5	94.6	993	3	US-10-010-577-1	Sequence 1, Appli
4	1680	93.4	3164	3	US-09-023-655-816	Sequence 816, App
5	1570	87.3	1126	3	US-10-010-580-1	Sequence 1, Appli
6	960.5	53.4	1482	2	US-08-330-121B-1	Sequence 1, Appli
7	960.5	53.4	1482	6	US-08-536-861-1	Sequence 1, Appli
8	960.5	53.4	1482	6	PCT-US95-13820-1	Sequence 1, Appli
9	960.5	53.4	1619	2	US-08-208-007A-1	Sequence 1, Appli
10	960.5	53.4	1619	3	US-08-915-095A-1	Sequence 1, Appli
11	960.5	53.4	1619	3	US-08-798-096-1	Sequence 1, Appli
12	960.5	53.4	1619	3	US-08-798-095A-1	Sequence 1, Appli
13	960.5	53.4	1619	3	US-09-953-956-1	Sequence 1, Appli
14	960.5	53.4	1619	3	US-08-553-125A-1	Sequence 1, Appli
15	960.5	53.4	1619	3	PCT-US94-04781-1	Sequence 1, Appli
16	960.5	53.4	1619	6	US-10-114-464-1	Sequence 1, Appli
17	960.5	53.4	1669	3	US-08-964-308-3	Sequence 3, Appli
18	960.5	53.4	1669	3	US-08-964-313-3	Sequence 3, Appli
19	960.5	53.4	1669	3	US-09-069-138-3	Sequence 3, Appli
20	956.5	53.2	1614	2	US-08-806-959-1	Sequence 1, Appli
21	953.5	53.0	990	2	US-08-684-932A-35	Sequence 35, Appl
22	951.5	52.9	990	3	US-08-964-308-13	Sequence 13, Appl
23	951.5	52.9	990	3	US-08-964-313-13	Sequence 13, Appl
24	951.5	52.9	990	3	US-09-069-138-13	Sequence 13, Appl
25	950.5	52.8	990	3	US-08-964-308-12	Sequence 12, Appl
26	950.5	52.8	990	3	US-08-964-313-12	Sequence 12, Appl
27	950.5	52.8	990	3	US-09-069-138-12	Sequence 12, Appl
28	862	47.9	1374	3	US-09-784-642-3	Sequence 3, Appli
29	859	47.7	1342	3	US-09-290-586A-2	Sequence 2, Appli
30	859	47.7	1342	3	US-09-290-586A-15	Sequence 15, Appl
31	859	47.7	1342	3	US-09-949-016-2733	Sequence 2733, Ap
32	848.5	47.2	1575	3	US-09-784-642-1	Sequence 1, Appli
33	848.5	47.2	1577	3	US-09-949-016-3909	Sequence 3909, Ap
34	848	47.1	1366	3	US-08-883-526-2	Sequence 2, Appli
35	848	47.1	1366	3	US-09-519-283-2	Sequence 2, Appli
36	803.5	44.7	1187	3	US-10-363-937-18	Sequence 18, Appl
37	796.5	44.3	1362	3	US-09-799-451-380	Sequence 380, App
38	756	42.0	1226	3	US-09-799-451-379	Sequence 379, App
39	726	40.4	1360	3	US-09-325-932A-98	Sequence 98, Appl
40	664.5	36.9	1600	3	US-09-325-932A-15	Sequence 15, Appl
41	650	36.1	1056	8	US-09-813-591-1	Sequence 1, Appli
42	650	36.1	1056	2	US-09-500-551-1	Sequence 1, Appli
43	641.5	35.7	651	2	US-08-546-712-1	Sequence 1, Appli
44	641.5	35.7	651	2	US-08-751-105-1	Sequence 1, Appli
45	636	35.4	1203	3	US-09-005-298-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-1227
Sequence 1227, Application US/09023655
Patent No. 6607879

GENERAL INFORMATION:
 APPLICANT: Cocks, Benjamin G.
 APPLICANT: Susan G. Stuart
 APPLICANT: Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 TITLE OF INVENTION: EXPRESSION
 NUMBER OF SEQUENCES: 1508
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 22, 2005, 09:43:40 ; Search time 3721 Seconds
                                             (without alignments)
                                             4161.925 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCLLLVCSSAVAQLHK..........MARNKGNHCGIASFPSYPEI 331

Scoring table: BLOSUM62
                Xgapop   10.0 , Xgapext 0.5
                Ygapop   10.0 , Ygapext 0.5
                Fgapop   6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
  -MODEL=frame+_p2n.model
  -Q=/Cgn2_l/USPTO_spool_p/US10809816/runat_21112005_094202_8297/app_query.fasta_1.519
  -DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
  -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
  -DOCALIGN=200 -THR_MAX=100 -THR_MIN=0 -ALIGN=10 -MODE=LOCAL
  -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
  -USER=US10809816@CGN_1_1_8010@runat_21112005_094202_8297 -LONGLOG
  -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -NCPU=6 -ICPU=3
  -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
  -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   EST:*
  1:   gb_est1:*
  2:   gb_est2:*
  3:   gb_est3:*
  4:   gb_htc:*
  5:   gb_est4:*
  6:   gb_est5:*
  7:   gb_est6:*
  8:   gb_est7:*
  9:   gb_est8:*
  10: gb_gss:*
  11: gb_gss2:*
       gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1747	97.1	1636	4	CR626534	full-leng
2	1747	97.1	1695	4	CR623699	full-leng
3	1747	97.1	1717	4	CR613968	full-leng
4	1747	97.1	1733	4	CR612707	full-leng
5	1552	86.3	896	4	AL548051	
6	1527	84.9	994	5	BX417536	
7	1519	84.4	953	1	AL513990	

No.						
8	1500.5	83.4	956	1	AL551671	AL551671
9	1486.5	82.6	1028	1	AL544870	AL544870
10	1406.5	78.2	1091	6	CF110629	CF110629	Shultzomi
11	1392	77.4	986	7	CO579243	CO579243	ILLUMIGEN
12	1386	77.0	1022	4	CR607339	CR607339	full-leng
13	1378.5	76.6	1355	4	AK028366	AK028366	Mus muscu
14	1345.5	74.8	944	8	DR772289	DR772289	ILLUMIGEN
15	1345.5	74.8	1287	4	BC011104	BC011104	Mus muscu
16	1333	74.1	928	8	DN824626	DN824626	ILLUMIGEN
17	1319.5	73.3	983	7	CO648649	CO648649	ILLUMIGEN
18	1319	73.3	945	7	CO646666	CO646666	ILLUMIGEN
19	1316	73.1	976	5	BX345726	BX345726	ILLUMIGEN
20	1304	72.5	848	8	BG539651	BG539651	602567596
21	1292	71.8	901	5	BW972127	BW972127	AGENCOURT
22	1292	71.8	904	8	DN824129	DN824129	LB0011.C2
23	1251	69.5	934	5	BQ710837	BQ710837	AGENCOURT
24	1247	69.3	908	5	BW975020	BW975020	AGENCOURT
25	1245	69.2	879	8	DN824771	DN824771	LB0011.C3
26	1238	68.8	876	8	DN824658	DN824658	LB0013.C2
27	1235	68.6	1068	3	BM806785	BM806785	AGENCOURT
28	1234	68.6	795	2	BG541942	BG541942	602569445
29	1231	68.4	839	1	AU140074	AU140074	602569445
30	1218	67.7	942	3	BP156785	BP156785	BP156785
31	1215	67.5	789	1	AM035440	AM035440	AM035440
32	1213	67.4	864	8	DN823937	DN823937	LB0011.C2
33	1209	67.2	764	8	DN997412	DN997412	TC117580
34	1209	67.2	855	8	DN824864	DN824864	LB0014.C2
35	1185	65.9	758	3	BI766075	BI766075	603053051
36	1181	65.6	759	7	CJ011620	CJ011620	CJ011620
37	1179	65.5	759	7	CK973194	CK973194	4104166 B
38	1175	65.3	782	2	BG548684	BG548684	602576991
39	1168	64.9	748	7	CK944131	CK944131	4068178 B
40	1161	64.5	809	6	CB993883	CB993883	AGENCOURT
41	1161	64.5	732	6	CB986537	CB986537	AGENCOURT
42	1160	64.5	710	6	CB986513	CB986513	AGENCOURT
43	1150	63.9	853	2	BE531143	BE531143	601278324
44	1145	63.6	702	8	DN272753	DN272753	LIB30321
45	1144	63.6	1048	2	BE616367	BE616367	601279178

ALIGNMENTS

RESULT 1
CR626534
LOCUS   CR626534   1636 bp   mRNA   linear   HTC 21-JUL-2004
DEFINITION   full-length cDNA clone CS0DI034YM16 of Placenta Cot 25-normalized
                of Homo sapiens (human).
ACCESSION   CR626534
VERSION   CR626534.1   GI:50507341
KEYWORDS   HTC; CNSLT_cDNA.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Hominidae; Homo.
REFERENCE   1 (bases 1 to 1636)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK   Contact : Peng Liang Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
                Faraday Avenue
REFERENCE   2 (bases 1 to 1636)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
  COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
                end enriched, double-strand cDNA was digested with Not I and cloned
                into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
                was normalized. Library was constructed by Life Technologies, a
                division of Invitrogen.

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 22, 2005, 08:23:59 ; Search time 573 Seconds
(without alignments)
3849.936 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCGIASPSYPEI 331

Searched: 4996997 segs, 333234308 residues

Scoring table: BLOSUM62
Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Fgapop 6.0 : Fgapext 7.0
Delop 6.0 : Delext 7.0

Total number of hits satisfying chosen parameters: 9993994

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/ccn21/USPTO_spool_p/US10809816/runat_21112005_094201_8278/app_query.fasta_1.519
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=10 -MODE=LOCAL
-OUTFMT=pto -NORMext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10809816 @CGN_1_1.1096 @runat_21112005_094201_8278 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1906s:*
2: geneseqn1990s:*
3: geneseqn2008s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

ALIGNMENTS

No.	Score	% Match	Length	DB	ID	Description
5	1747	97.1	4100	14	ADY14943	DNA encod
6	1747	97.1	4100	14	AEC01586	Human cat
7	1747	97.1	1254	12	ADJ71694	Human NOV
8	1744	96.9	1752	11	ADP65424	Human, ca
9	1744	96.9	1643	3	AAZ56150	Human Cat
10	1740	96.7	1763	6	ABL61836	Colon ade
11	1740	96.7	1763	11	ABK84111	Human cDN
12	1740	96.7	1763	11	ADP65426	Human cat
13	1738	96.6	1255	10	ADE84859	Farnesyl
14	1738	96.6	1255	10	ADR25126	Breast ca
15	1738	96.6	1255	14	ADY19783	DNA encod
16	1701.5	94.6	993	9	AAL61281	Monkey ca
17	1680	93.4	3164	13	ADI31490	Human cDN
18	1680	93.4	3164	13	ADS83557	Human lym
19	1631	90.7	1552	13	ACN41976	Human dia
20	1631	90.7	1794	13	ACN41975	Human dia
21	1600	88.9	996	8	ACC42956	Canine ca
22	1580	87.8	996	8	ABX12290	Human Cat
23	1577	87.7	1895	10	ADG75772	Human pro
24	1558	86.6	996	8	ABX15396	Human Cat
25	1330	73.3	996	8	ADB57919	Toxicity-
26	1330	73.3	1330	10	ADB52867	Primary r
27	1318.5	73.3	1330	10	ABT41998	Toxicity
28	1318.5	73.3	1330	12	ADP72587	Renal tox
29	1318.5	73.3	1330	13	ADV41136	Rat cardi
30	1318.5	73.3	1330	13	ADJ71696	Rat cardi
31	1144	73.3	667	12	ADZ11833	Canine ca
32	963	53.5	1001	14	AAC62064	DNA encod
33	961.5	53.4	990	3	AAV43859	Human cat
34	960.5	53.4	990	3	AAT27040	Prepro-ca
35	960.5	53.4	1482	2	ADQ18391	Human sof
36	960.5	53.4	1482	12	AAT04002	Breast ca
37	960.5	53.4	1619	2	ADR24618	Breast ca
38	960.5	53.4	1669	13	ACN37459	Tumour-as
39	960.5	53.4	1669	13	ADP25399	PRO polyp
40	960.5	53.4	1669	13	ABX63462	Human cDN
41	960.5	53.4	1671	8	ACA64820	Human cat
42	960.5	53.4	1671	8	ADP12815	Reference
43	960.5	53.4	1702	14	AEC01584	Human cat
44	960.5	53.4	1702	14	AAF16318	Human pro
45	960.5	53.4	2017	11	ACN89036	Breast ca

RESULT 1
ADI31901

ID   ADI31901 standard; cDNA; 1784 BP.
XX
AC   ADI31901;
XX
DT   17-JUN-2004 (first entry)
XX
DE   Human cDNA #1227.
XX
KW   Human; gene; ss; immunological response; immunopathological condition;
KW   Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW   irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW   acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
KW   osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US6607879-B1.
XX
PD   19-AUG-2003.
XX
PF   09-FEB-1998; 98US-00023655.
XX
PR   09-FEB-1998; 98US-00023655.
XX
PA   (INCY-) INCYTE CORP.
XX

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1747	97.1	1784	11	ADI31901	Human cDN
2	1747	97.1	1784	13	ADS83968	Human lym
3	1747	97.1	4100	11	ADP64970	Human cat
4	1747	97.1	4100	14	ADY14945	DNA encod

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 22, 2005, 09:26:44 ; Search time 4594 Seconds
                 (without alignments)
                 4095.599 Million cell updates/sec

Title:
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK...........MARNKGNHCGIASPPSYPEI 331

Scoring table: BLOSUM62
                 Xgapop 10.0 , Xgapext 0.5
                 Ygapop 10.0 , Ygapext 0.5
                 Fgapop   6.0 , Fgapext 7.0
                 Delop   6.0 , Delext   7.0

Searched: 5883141 seqs, 2842172565653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10809816/runat_21112005_094201_8286/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=10 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10809816 @CGN 1 1.7415 @runat_21112005 094201_8286 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
            1: gb_ba:*
            2: gb_in:*
            3: gb_env:*
            4: gb_om:*
            5: gb_ov:*
            6: gb_pat:*
            7: gb_ph:*
            8: gb_pr:*
            9: gb_ro:*
           10: gb_sts:*
           11: gb_un:*
           12: gb_gy:*
           13: gb_vi:*
           14: gb_heg:*
           15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	Match	ID	Description
1	1747	97.1	996	11		AY888599	Synthetic
2	1747	97.1	996	11		AY891245	Synthetic
3	1747	97.1	996	11		AY893335	Synthetic

4	1747	97.1	1259	6	CQ728273	CQ728273	Sequence
5	1747	97.1	1784	6	AR380682	AR380682	Sequence
6	1747	97.1	4100	6	CS031243	CS031243	Sequence
7	1747	97.1	4100	6	CS031245	CS031245	Sequence
8	1747	97.1	1763	8	CS040195	CS040195	Sequence
9	1747	97.1	4100	6	CS040197	CS040197	Sequence
10	1747	97.1	4100	6	CS118410	CS118410	Sequence
11	1747	96.9	996	8	CR541676	CR541676	Homo sapi
12	1744	96.9	996	11	AY893542	AY893542	Synthetic
13	1744	96.9	1743	8	BC002642	BC002642	Homo sapi
14	1740	96.7	1643	6	BD245063	BD245063	Use of ca
15	1740	96.7	1763	6	AR209918	AR209918	Sequence
16	1740	96.7	1763	8	AX329664	AX329664	Sequence
17	1738	96.6	1255	8	HUMCATS	HUMCATS	Homo sapien
18	1738	96.6	1255	8	M90696	M90696	Homo sapien
19	1738	96.6	1255	6	CS036083	CS036083	Sequence
20	1738	96.6	1255	6	CS045035	CS045035	Sequence
21	1738	96.6	1255	6	AX774762	AX774762	Sequence
22	1738	96.6	1255	8	S93414	S93414	cathepsin S
23	1738	96.6	993	6	AR579499	AR579499	Sequence
24	1701.5	94.6	993	6	AX774253	AX774253	Sequence
25	1701.5	94.6	1143	8	AY156691	AY156691	Saimiri b
26	1680	93.4	3164	6	AR380271	AR380271	Sequence
27	1600	88.9	996	4	AX797393	AX797393	Sequence
28	1600	88.9	1125	4	AY156692	AY156692	Canis fam
29	1570	87.3	1126	6	AR568030	AR568030	Sequence
30	1406.5	78.2	1391	6	BC059142	BC059142	Rattus no
31	1366.5	76.0	1279	9	BC002125	BC002125	Mus muscu
32	1361.5	75.7	1279	9	AF038546	AF038546	Mus muscu
33	1349.5	75.0	1296	9	MMCATHEPS	MMCATHEPS	Mus muscu
34	1318.5	73.3	1330	9	AJ002386	AJ002386	Mus muscu
35	1230.5	68.4	1499	5	L03201	L03201	Rattus norv
36	1140.5	63.4	1752	5	AJ719318	AJ719318	Gallus ga
37	1140	63.4	1778	5	BC077285	BC077285	Xenopus l
38	1112.5	61.8	2248	5	BC056059	BC056059	Xenopus t
39	1094	60.8	1088	5	AY333300	AY333300	Paralabid
40	1070	59.5	1242	5	BC055788	BC055788	Danio rer
41	1065	59.2	1246	5	BC095694	BC095694	Danio rer
c 42	1053	58.5	601	10	BV179264	BV179264	sqnm10459
43	1050	58.4	1238	5	BC093339	BC093339	Danio rer
44	1047	58.2	1257	5	BC096862	BC096862	Danio rer
45	1043	58.0	1399	5	AY622858	AY622858	Oncorhync

ALIGNMENTS

RESULT 1
AY888599
LOCUS   AY888599   996 bp   mRNA   linear   SYN 22-MAR-2005
DEFINITION   Synthetic construct Homo sapiens clone FLH019213.01X cathepsin S
                (CTSS) mRNA, complete cds.
ACCESSION   AY888599
VERSION   AY888599.1 GI:61358281
KEYWORDS   Human ORF Project.
SOURCE   synthetic construct
  ORGANISM   synthetic construct
                other sequences; artificial sequences.
REFERENCE   1 (bases 1 to 996)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
                Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
                Williamson,J. and LaBaer,J.
  TITLE   Cloning of human full-length CDS in Creator (TM) recombinational
                vector system
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 996)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
                Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
                Williamson,J. and LaBaer,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-JAN-2005) Biological Chemistry and Molecular
                Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
                Cambridge, MA 02141, USA
COMMENT   This CDS clone is a part of a collection of human full-length

us-10-809-816a-2.rapbm

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 21, 2005, 12:48:40 ; Search time 164 Seconds
(without alignments)
843.302 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK...........MARNKGNHCGIASFPSYPEI 331
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1799	100.0	331	5	US-10-809-816A-2	Sequence 2, Appli
2	1747	97.1	331	5	US-10-973-858-20	Sequence 20, Appli
3	1747	97.1	331	5	US-10-287-4364-412	Sequence 412, App
4	1747	97.1	331	5	US-10-287-4364-1109	Sequence 1109, Ap
5	1744	96.9	331	5	US-10-635-398-96	Sequence 96, Appl
6	1744	96.9	331	5	US-10-809-816A-4	Sequence 4, Appl
7	1744	96.9	331	5	US-10-273-577-1	Sequence 1, Appli
8	1741	96.7	331	4	US-10-099-275-2	Sequence 2, Appli
9	1740	96.6	331	3	US-09-953-956-8	Sequence 8, Appli
10	1738	96.6	331	4	US-10-114-464-8	Sequence 8, Appli
11	1738	96.6	331	4	US-10-726-645-8	Sequence 8, Appli
12	1734	96.4	331	4	US-10-273-577-6	Sequence 6, Appli
13	1731	96.2	331	4	US-10-318-584-4	Sequence 4, Appli
14	1731	96.2	331	4	US-10-318-584-4	Sequence 4, Appli
15	1724	95.8	331	4	US-10-273-577-7	Sequence 7, Appli
16	1724	95.8	330	5	US-10-010-577-2	Sequence 2, Appli
17	1701.5	94.6	330	4	US-10-894-104-2	Sequence 2, Appli
18	1701.5	94.6	330	5	US-10-809-816A-3	Sequence 3, Appli
19	1645	91.4	322	5	US-10-809-816A-7	Sequence 7, Appli
20	1600	88.9	331	4	US-10-010-580-2	Sequence 2, Appli
21	1600	88.9	331	4	US-10-894-106-2	Sequence 2, Appli
22	1600	88.9	331	5	US-10-809-816A-5	Sequence 5, Appli
23	1580	87.8	331	3	US-10-990-064-2	Sequence 2, Appli
24	1388	77.2	267	4	US-10-408-765A-1343	Sequence 1343, Ap
25	1338.5	74.4	330	4	US-10-809-816A-6	Sequence 6, Appli
26	1145	63.6	217	4	US-10-273-577-5	Sequence 5, Appli
27	1144	63.6	222	5	US-10-635-398-98	Sequence 98, Appl
28	1138	63.3	217	4	US-10-273-577-2	Sequence 2, Appli
29	1135	63.1	217	4	US-10-273-577-4	Sequence 4, Appli
30	1128	62.7	217	4	US-10-273-577-3	Sequence 3, Appli
31	965.5	53.7	329	4	US-09-953-956-7	Sequence 7, Appli
32	965.5	53.7	329	4	US-10-318-584-7	Sequence 7, Appli
33	965.5	53.7	329	5	US-10-809-816A-1	Sequence 1, Appli
34	965	53.6	329	5	US-10-726-645-7	Sequence 7, Appli
35	963	53.5	330	5	US-10-929-919A-32	Sequence 32, Appl
36	960.5	53.4	329	3	US-09-953-956-2	Sequence 2, Appli
37	960.5	53.4	329	4	US-10-114-464-2	Sequence 2, Appli
38	960.5	53.4	329	5	US-10-318-584-2	Sequence 2, Appli
39	960.5	53.4	329	4	US-10-726-645-2	Sequence 2, Appli
40	960.5	53.4	329	5	US-10-723-860-1211	Sequence 1211, Ap
41	960.5	53.4	329	5	US-10-973-858-18	Sequence 18, Appl
42	960.5	53.4	329	5	US-10-756-149-5007	Sequence 5007, Ap
43	960.5	53.4	361	3	US-09-925-300-1693	Sequence 1693, Ap
44	877	48.7	291	4	US-10-836-897-4	Sequence 4, Appli
45	862	47.9	334	3	US-09-784-641-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-809-816A-2

; Sequence 2, Application US/10809816A
; Publication No. US20050214774A1
; GENERAL INFORMATION:
; APPLICANT: LI, Shyr-Jiann et al.
; TITLE OF INVENTION: ISOLATED MONKEY CATHEPSIN S PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING MONKEY CATHEPSIN S PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001507
; CURRENT APPLICATION NUMBER: US/10/809,816A
; CURRENT FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Cynomologous monkey
US-10-809-816A-2

Query Match 100.0%;
Best Local Similarity 100.0%;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Score 1799; DB 5; Length 331;
Pred. No. 5.6e-165;

Qy   1 MKQLVCVLLVCSSAVAQLHKDPTLDHHWHLWKKTYGKQYKEKNEEAVRRLIWEKNLKFVM 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKQLVCVLLVCSSAVAQLHKDPTLDHHWHLWKKTYGKQYKEKNEEAVRRLIWEKNLKFVM 60

Qy   61 LHNLEHSMGMHSYDLGMNHLGDMTSEEVMSLMSSLRVPSQWQRNITYKSNANQILPDSVD 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LHNLEHSMGMHSYDLGMNHLGDMTSEEVMSLMSSLRVPSQWQRNITYKSNANQILPDSVD 120

Qy 121 WREKGCVTEVKRYQGSCGACWAFSAVGALEAQLKIKTGKLVSLSAQNLVDCSTEKYGNKGC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 WREKGCVTEVKRYQGSCGACWAFSAVGALEAQLKLKTGKLVSLSAQNLVDCSTEKYGNKGC 180

Qy 181 NGGFMTRAFQYIIDNNGIGDSDASYPYKATDQKCQYDSKYRAATCSKYTELPYGREDVLKE 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 NGGFMTRAFQYIIDNNGIGDSDASYPYKATDQKCQYDSKYRAATCSKYTELPYGREDVLKE 240

Qy 241 VVANKGPVSVGVGDASHPSFFLYRSGVYYEPSCTQNVNHGVLVVGYGVLNGKEYWLVKNSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 VVANKGPVSVGVGDASHPSFFLYRSGVYYEPSCTQNVNHGVLVVGYGVLNGKEYWLVKNSW 300

Qy 301 GRNFGEEGYIRMARNKGNHCGIASFPSYPEI 331
          ||||||||||||||||||||||||||||||
Db 301 GRNFGEEGYIRMARNKGNHCGIASFPSYPEI 331

RESULT 2

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: November 21, 2005, 12:41:29 ; Search time 24 Seconds
(without alignments)
1140.235 Million cell updates/sec

Title: US-10-809-816A-2
Perfect score: 1799
Sequence: 1 MKQLVCVLLVCSSAVAQLHK..........MARNKGNHCGIASPSYPEI 331
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptcdata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptcdata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptcdata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptcdata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptcdata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptcdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1740	96.7	331	1	US-08-208-007A-8	Sequence 8, Appli
2	1740	96.7	331	1	US-09-701-685-2	Sequence 2, Appl
3	1740	96.7	331	2	US-09-290-586A-20	Sequence 20, Appl
4	1738	96.6	331	2	US-08-860-255A-2	Sequence 2, Appli
5	1738	96.6	331	2	US-08-915-095A-8	Sequence 8, Appli
6	1738	96.6	331	2	US-08-798-096-8	Sequence 8, Appli
7	1738	96.6	331	2	US-08-798-095A-8	Sequence 8, Appli
8	1738	96.6	331	2	US-09-953-956-8	Sequence 8, Appli
9	1738	96.6	331	2	US-08-553-125A-8	Sequence 8, Appli
10	1738	96.6	331	2	US-10-114-464-8	Sequence 8, Appli
11	1731	96.2	331	2	US-08-330-121B-4	Sequence 4, Appli
12	1731	96.2	331	4	PCT-US95-861-4	Sequence 4, Appli
13	1731	96.2	331	4	PCT-US95-13820-4	Sequence 4, Appli
14	1701.5	94.6	330	2	US-08-208-007A-7	Sequence 2, Appli
15	1600	88.9	331	1	US-10-010-580-2	Sequence 2, Appli
16	1600	88.9	329	1	US-10-010-577-2	Sequence 2, Appli
17	965.5	53.7	329	2	US-08-330-121B-3	Sequence 3, Appli
18	965.5	53.7	329	2	US-08-915-095A-7	Sequence 7, Appli
19	965.5	53.7	329	2	US-08-798-096-7	Sequence 7, Appli
20	965.5	53.7	329	2	US-08-798-095A-7	Sequence 7, Appli
21	965.5	53.7	329	2	US-09-953-956-7	Sequence 7, Appli
22	965.5	53.7	329	2	US-08-553-125A-7	Sequence 7, Appli
23	965.5	53.7	329	2	US-08-536-861-3	Sequence 3, Appli
24	965.5	53.7	329	2	US-10-114-464-7	Sequence 7, Appli
25	965.5	53.7	329	4	PCT-US95-13820-3	Sequence 3, Appli
26	960.5	53.4	329	1	US-08-330-121B-2	Sequence 2, Appli
27	960.5	53.4	329	1	US-08-852-807-20	Sequence 20, Appl
28	960.5	53.4	329	2	US-08-964-308-4	Sequence 4, Appli
29	960.5	53.4	329	2	US-08-964-313-4	Sequence 4, Appli
30	960.5	53.4	329	2	US-08-860-255A-1	Sequence 2, Appli
31	960.5	53.4	329	2	US-09-069-138-4	Sequence 4, Appli
32	960.5	53.4	329	2	US-08-915-095A-2	Sequence 2, Appli
33	960.5	53.4	329	2	US-08-798-096-2	Sequence 2, Appli
34	960.5	53.4	329	2	US-08-798-095A-2	Sequence 2, Appli
35	960.5	53.4	329	2	US-09-953-956-2	Sequence 2, Appli
36	960.5	53.4	329	2	US-08-553-125A-2	Sequence 2, Appli
37	960.5	53.4	329	2	US-10-114-464-7	Sequence 2, Appli
38	960.5	53.4	329	2	US-08-536-861-2	Sequence 2, Appli
39	960.5	53.4	329	2	US-09-290-586A-19	Sequence 19, Appl
40	960.5	53.2	329	4	PCT-US95-13820-2	Sequence 2, Appli
41	956.5	53.2	329	1	US-08-806-959-2	Sequence 2, Appli
42	954.5	53.1	329	2	US-08-208-007A-2	Sequence 2, Appli
43	954.5	53.1	329	4	PCT-US94-04781-2	Sequence 2, Appli
44	953.5	53.0	329	2	US-08-684-932A-36	Sequence 36, Appl
45	951.5	52.9	329	2	US-08-964-308-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-208-007A-8
; Sequence 8, Application US/08208007A
; Patent No. 5501969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, ET AL.
; TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,007A
; FILING DATE: March 8, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: No. 5501969e
; FILING DATE: No. 5501969e
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-208-007A-8

Query Match 96.7%; Score 1740; DB 1; Length 331;
Best Local Similarity 96.7%; Pred. No. 2.9e-173;
Matches 320; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKQLVCVLLVCSSAVAQLHKDPTLDHHHHLWKKTYGKOYKEKNEEAVRRLIWEKNLKFVM 60
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||